Data Analysis and Decision Making - II
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 47
Goal programming

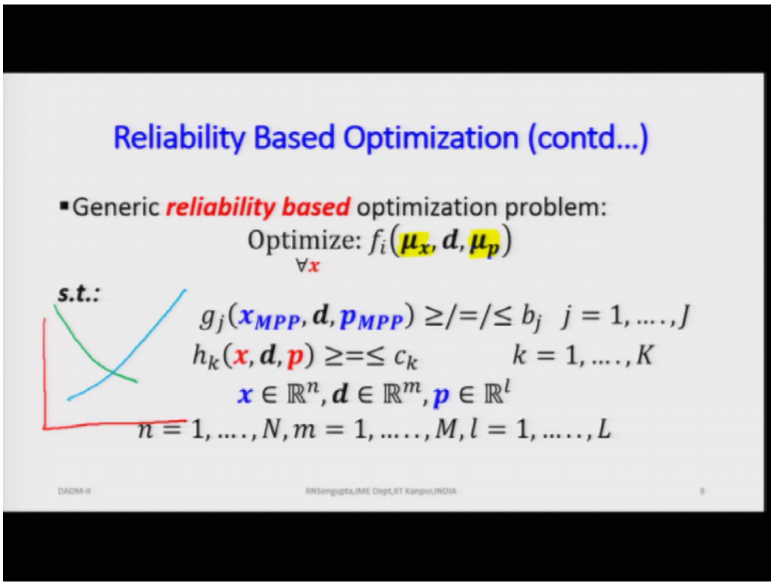
Welcome back my dear friends; a very good morning, good afternoon and good evening to all of you wherever you are and as you know this is the DADM II which is Data Analysis and Decision Making II course under the NPTEL MOOC series and this total course duration is 12 weeks which is 30 hours which when converted into number of lectures this 60 number because, each lecture is for half an hour. And as you can see in the slide we are in the 47th lecture which is in the 10th week and after each week which is 5 lectures each being for half an hour you have 1 assignment and I am sure you have done pretty well in your first 9 assignments.

So, we are covering the concept of reliability, optimization and some concepts. So, rather than solving problems it will be more very simple definitions related assignments because this is quite involved. So, I will go slowly in order to explain you the general concepts and my good name is Raghu Nandan Sengupta from the IME department at IIT Kanpur.

So, if you remember we at considering the PMA and RI approach where it the concept was you had in x space which room 1, u space in room number 2, u am just giving you the bullet points. So, whatever the problem solution is or the problem given is, you first have the deterministic problem and consider the inputs based on the fact they are mean values, solve them you get some optimum values.

So, you consider those optimum values which technically should be the mean values, using those mean values you basically project them in the u space, find out the u values for x and p, use either the PMA or the RI approach where either the disk increases its size and till reaches that boundary of the feasible region on infeasible region whatever you say. In another case that disk remain the same depending on a normalization value of beta. The boundary moves such that in both the cases the moment it touches you get the first optimum solution. Map it back to u space, again solve the iteration the optimization problem using the methodology which you know, find out the next set of x's and p's. So, x's are the values which you want to want decision variables, again map it back to the u space, continue doing it till the difference in the objective function values f of x is bounded by an epsilon value.

(Refer Slide Time: 02:58)



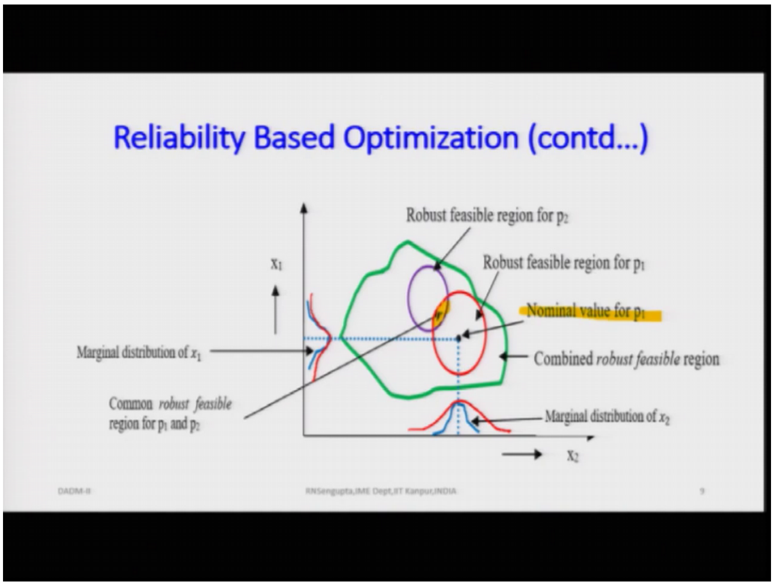
So, now how you do it? So, in generic reliability based optimization problem you have that functional value f i where, i is equal to 1, 2 whatever values of capital I you have and this mean values which you have the mean values which are blue in colour which I am now highlighting with yellow, this p values are the parameter values which are external and this x are the decision variable. So, the mean values mean they are the mean.

So, I will just draw it very simply. So, this was basically I will use only one colour. So, it is easy for to explain. So, this is boundary one of the boundaries this is another at the boundaries. So, these are the mean values the blue one and the green one are the mean values of x which x based on which you are trying to solve. So now, once you find out the value. So, using the RI and the PMA method you are trying to find out the most probable point MPP for x and p based on that you will solve. So, d would not have because d is basically the deterministic value.

So, once you solve it, you will find out the most probable point and the corresponding x value. So, remember one thing when I am trying to solve it I am trying to make a differentiation between the blue colour and the red colour. Red colour I will consider the probabilistic part, the count and the counterpart in the blue would be the deterministic part. But the issue is this x MPP is basically the most probable point in the u space while is counterpart here would be the mu x and the values as it changes per iteration value.

So, you have the constraints depending on the probabilistic constraints and the deterministic constraints and obviously you know x is along the real line in dimension, d is along the again real values are in the m dimension and p is in the l dimension depending on the values of m n and l.

(Refer Slide Time: 05:13)



Now, if I consider from the diagrammatic point of view. So, you are basically trying to solve the problem both from the reliability part as well as the robust part. So, robust you are trying to ensure that it will give you some very good solutions as that it perturbation values being there, it the overall solution does not change much. So, they would be some the some variances, but the variability would be as low as possible let me put it in this way.

So, what you want to solve is, is that you want to find out the areas based on the robust framework area and the reliability part. So, what you are trying to do again I am considering the distributions in the very simplistic case to be normal. So, we will consider the robust feasible region for p 2; p 2 is basically the probabilistic external variable which is there. So, they can be p 1, p 2, p 3 and the vector is p. Similarly you will have basically the robust feasible region for p 1 also. So, p 1 p 2 would give you intersection which is hashed. Here would give you the intersection based on which you will try to solve the problem. So, that will give you the common robust feasible region for p 1 p 2. So, if it is a higher dimension, it will be a intersection with three-dimension it will be intersection of 2 spheres or 3 spheres and in the higher cases it will be a intersection of the hyper spheres.

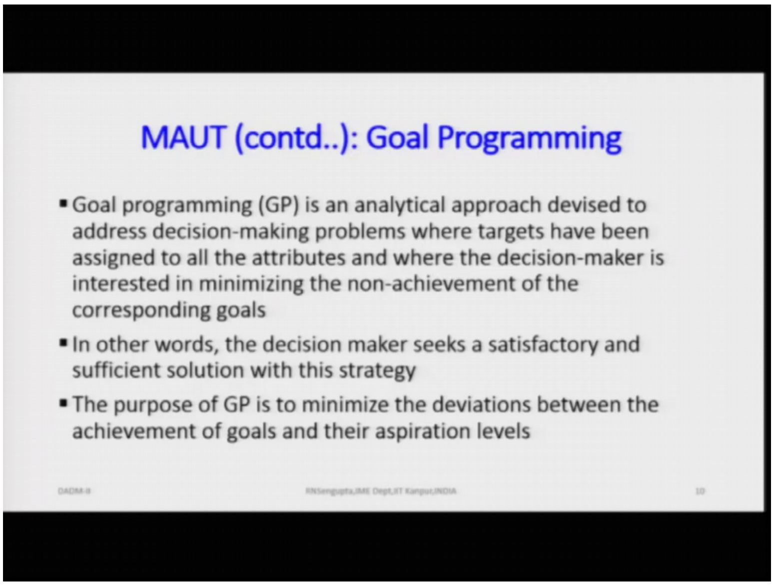
Now, when you find out the nominal value, nominal value is basically some mean value based on which the perturbation or the vibration is happening considering in the very simple way vibration of an atom. So, as the vibration happens the overall area of influence would basically be dictated on the level of energy when you if you know in very simple physics or chemistry, but what I am trying to do is that higher the vibration would basically mean higher, the higher the perturbation is. Lower it is, less the perturbation is. That means, the variability is low and in that case variability is high. That means, in the case where the variability is low, your reliable solution is much more robust and much better in the sense that you can save with certainty, that is less variability.

So, you have a nominal value for p 1. Similarly we will have nominal value of p 2 also and the combined area of the robust region because, why I am saying the combined region is would basically it will depend on different able constraints. So, the green area which I have drawn is basically the combined area and the violet one and the red one are basically based on the perturbation of p 1 and p 2.

Now, remember one thing when you have the perturbations of equal values and you consider the simple concept of normal distribution or consider that you are considering the concept of central limit theory to be true and you have bring in to the picture the concept of normal distribution, then the distributions which you have let me draw it. Then the distributions which you have the univariate case for x or p and x and can be x 1, x 2, x 3 or p can be p 1, p p 3. It will be a normal distribution depending on the variability, again this is a normal distribution and as already explained variance being same it will be a circle variable variability being different, it will be a rugby ball depending on how high or low the variability is along x 1 and x 2 direction considering you have x 1 and x 2 only.

Now it may so happen, now as I said I am repeating it again that the variability being same and variability being different whatever the case if it is normal then solving it using the standard normal deviation deviate problem in the multivariate case is also simple. The moment it is non-normal then trying to solve it using simulation method is the only way. So, you basically simulate it and try to find it out.

(Refer Slide Time: 09:29)



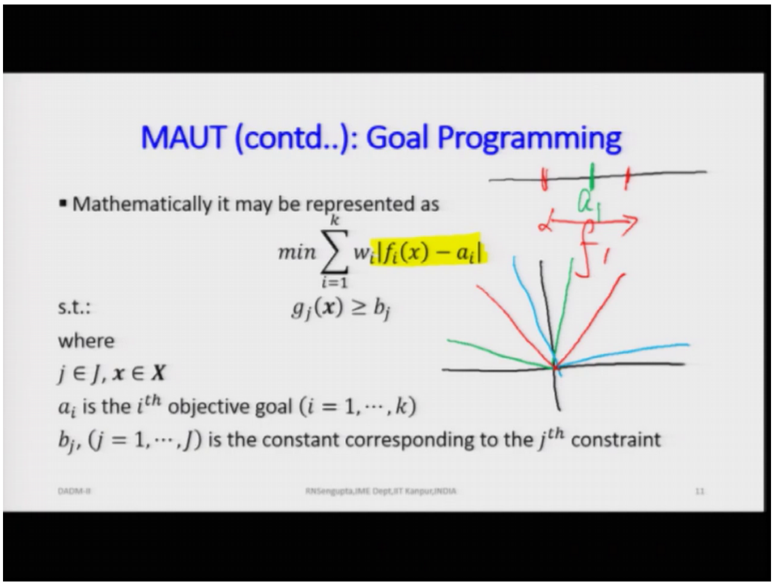
Now, we will consider very simple concept of multi attribute utility theory considering very simple concept of Goal Programming. So, by the way these solution methodologies I am not going to come. I am only giving you the simple background of robust optimization as well as goal programming. So, when goal programming this basically analytical approach devised to address decision making problems where targets have been assigned to all the attributes.

So, targets are there attribute values are given it can be high and low and you want to achieve it some attribute value depending on what you think is the best. So, consider there is some concept of risk. Now the concept of risk I am utilizing in such a sense that my level of satisfaction for attaining some goal would be definitely different from your level of risk or your level of satisfaction to at attaining that goal. So, I am using the concept of risk or the level of reliability or probability whatever you say. So, where the targets have been assigned to all the attributes and where the decision maker is interested in minimizing the non-achievement of the corresponding goals.

So, say for example, my goal is basically to achieve a an attribute value for whatever decision it is consider it is say for example, on a scale of 1 to 10 is 7. So, obviously they would be variability. So, I want to basically minimize the overall variability of not attaining my goal. So, in other words the decision maker seeks the satisfactory and sufficient solution with his or her strategy such that with minimum number of probability or chance that attribute would not be attained. Obviously, you cannot make it 0. So, if you think that each and every time that attribute level will be attain attained, that is not true because there is a perturbation there is a reliability, there is a probability or there is a dispersion.

So, the purpose of gaol programming is to minimize the deviations between the achievement of the goals and their aspiration values based on which you can achieve the best set of results. So, how would you do it? So, mathematically if you want to basically minimize the perturbation what you will do is that you want to basically minimize the deviation between the functional value and the attainable value or the attribute which you have.

(Refer Slide Time: 11:48)



So, what you will do is that, for each value of functional value of f 1 you have an attribute a 1 and you want to basically minimize the dispersion of f 1 and a 1; similarly from f 2 a 2 and so on and so forth. Now I am not mentioning f 1 x or f 2 x because x is basically vector that can be done accordingly. So, I am only considering the functional form. So, what I want to do so, consider that actual value obtained that mean the black line, this is a 1 value and there is some perturbations based on the fact that f 1 can be between some a 1 plus delta in a 1 minus delta and obviously the question would come that what if the perturbations are unequally penalized.

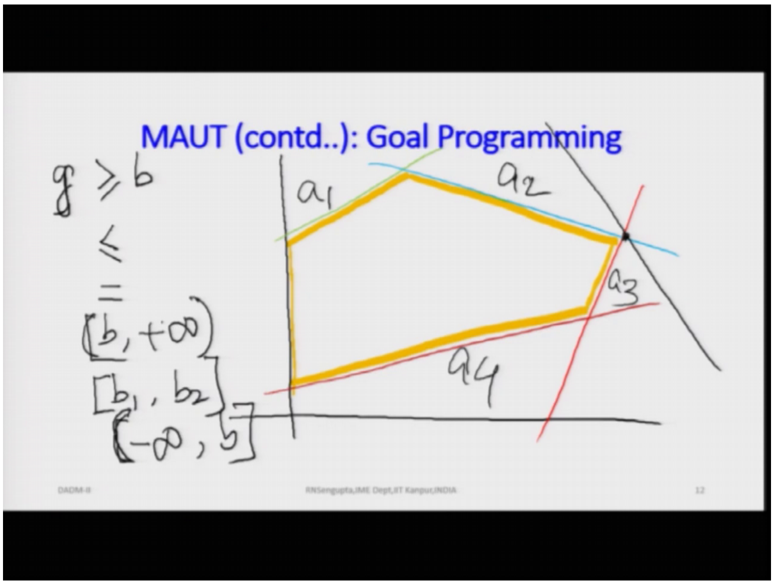
Now if you remember I did discussed about some penalty functions and the linex laws and the effect of that. So, again I will repeat those three examples. One was for the on the dam building the dam. Height was actually should be 120, but you have building it to either 118 or 122; so, in that case when if it is 118, the overall loss in the case when it was negative was much higher because, for the catastrophic loss. In the case if it is over estimated then the level of loss would be less.

So obviously, the perturbations on two different directions would be different. In the case when you have the case for electrical circuit the big electrical circuit which you have and the vacuum circuit breakers are there. So, if you over estimate the overall loss and if you basically replace those products later on after the guaranteed life on an average then the pert the overall loss would be much higher because the chance on accident is much higher. In the case when it is under estimate, then you replace that products more frequently. So, overall loss would be minimized because you only stoppages and manual loss would be there not any a catastrophic loss.

Similarly, you can formulate another problem where over estimation or an estimation would be unequally equally penalized depending on the type of set of the problem which you have. Now, what you are trying to do is that you are trying to give weight ages. So, the weight ages and again if you remember for the for the unequal penalty loss if it was a linlin loss function. So, overall value, so, if you have this I will use the same colour for the functions. So, this under utilization is less penalized, over utilization is more penalized in this case, under utilization is more penalized over estimation is less penalized and in this case under estimation is equally penalized as over estimation.

So, if you have I am sure you have seen these diagrams when I have drawn. So, we will give and they can be weighted. Weighted means you are giving some weights. So, these would be w 1, w 3, w 1, w 2, w 3 depending on number of such the goals you have to attain. So, a i is the ith objective function which you want to attain on the attribute and b j is the constant depending on the jth constraint values which you have and you want to basically have the constraints accordingly such the g 1 is greater than equal to b 1, g 2 is greater than b 2and so on and so forth.

(Refer Slide Time: 15:48)



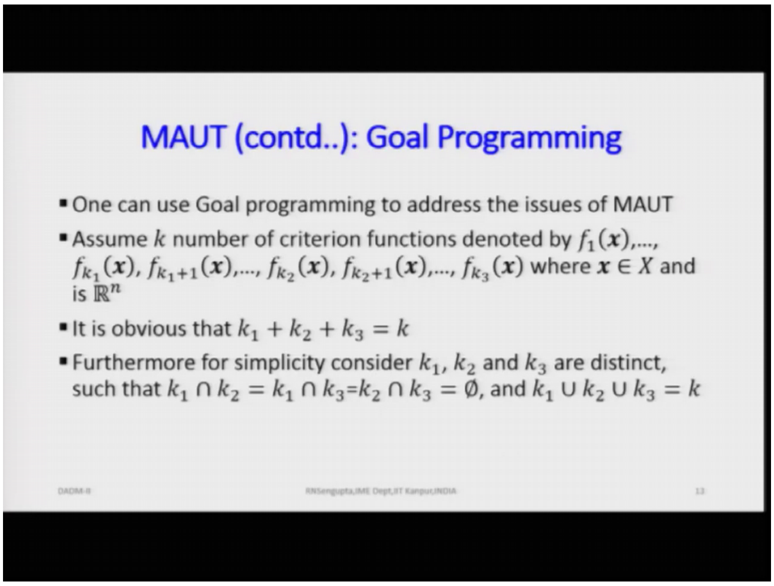
So, what you want to have, so this is so, you have the constraints would be coming now. So, constraints this is one of the constraints. I am drawing the simple linear constraints. The second one is this, then third one is this, then fourth one is this. So, now you are so, let me use another want wait it will be easy for me. So, I will use this and I will use the objective function with the black line so, because the constraints have been given with different colours. So, this is the, this drawing the boundary. So, this is the feasible region and what you have the constraints. So, constraint basically is in maximization one it becomes maximum at say for example, this is point I am trying to draw it as neat as possible.

Now, what I considered is that the attainable points are the constraints are these fixed lines which is a 1, a 2, a 3, a 4. So, I mark them as a 1, a 2, a 3, a 4 and I will give weights accordingly to those constraints depending on the level of importance. So, if I give the if the perturbation is very low and the weights is also low that that means I would not be so much worried about the constraints being violate or non-violative because I have already considered in the objective function. What I have done in the objective function is basically to find out the difference between the attainable values, by the way these let me see these values are not the a's which I have for the objective function.

So, these are I am just mentioning them as a 1, a 2, a 3, a 4, but the actual functional form which was f 1 minus a 1 mod of that, f 2 minus a 2 mod of that are that attainable value which you want to do. So, say for example, for the objective function which you want to maximize you think for the cost; for the cost you will minimize for say for example, if revenues you want to maximize you want to attain some level of values. Say for example, you want to achieve attain and a profit per month of 2,00,00,000.

So, that is the a value and based on that what is the, the weight you will get. So, if you for you if you think that your level of attainment has to be done as my minutely as possible then the weights you will put accordingly would be high or low depending on what importance do you place on that objective function.

(Refer Slide Time: 18:58)

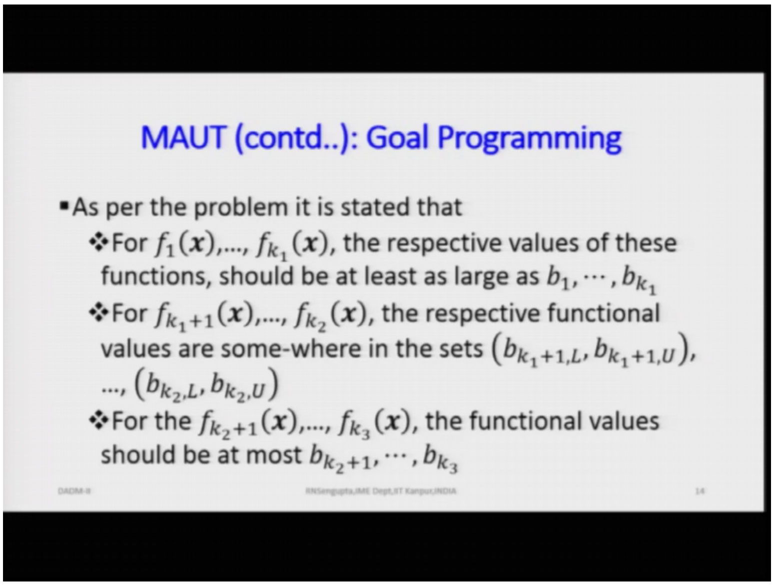


So, one can use goal programming to address the issues of multi attribute utility theory. So, let us consider in this way assume k number of criteria functions denote denoted by f 1 to f k and you basically club them into different groups. So, you have basically f 1 to k 1 then k 1 plus 1 to f k 2. So, whole set of k is being divided another set would be k 2 plus 1 to k 3 such that it is obviously true that the union or the some of k 1 k 2 k 3 is k.

So, you have basically divided the whole set of objective functions into three groups, k 1, k 2, k 3 and you will also consider why you are doing it because you want to basically achieve some maximization some minimization for them. So, if k 1, k 2 are on are only sets. So, you will basically have the maximization sets as k 1, the minimization sets are k

2 and we will consider obviously they would not be any intersections set between k 1, k 2, they would be null set.

(Refer Slide Time: 20:08)



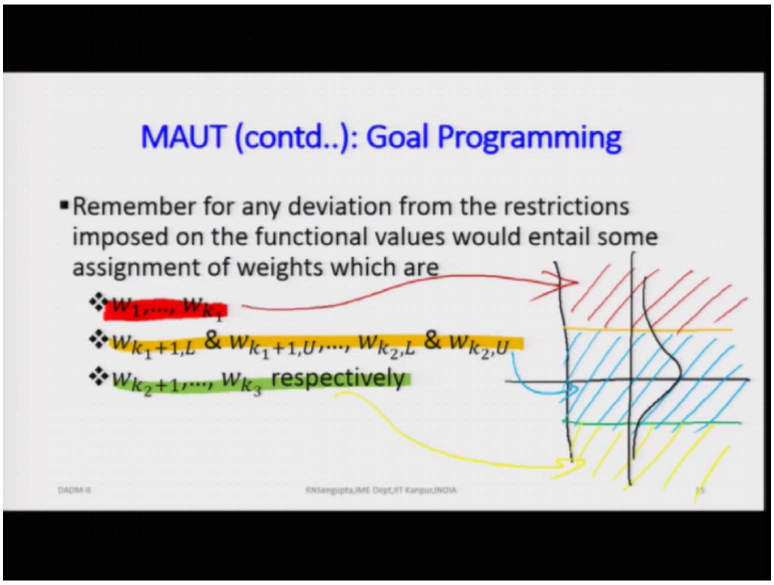
Now, as per the problem what is stated is that you will basically at want to attain f 1 to f k 1, the respective values of these functions should be at least as large as b 1 to b k 1. For the other case you will basically want to have f k 1 plus 1 to f k 2 at the respective functional forms. Somewhere in that set between the bounds which you have put for yourself for the k plus 1 till k 2 sets of those objective functions. Because, what you have in the in the greater than side sign you have b 1 to b k 1, then b k 1 b k 1 plus 1 till b k 2 and b then b. These k's I am mentioning in the suffix. So, it will be b k 2 plus 1 till b k.

So, you have want to achieve in different proportions on different values. For the b 1 sets, the first set you want to be find out them as large as possible, in another set they would be some bound between which you want to attain and for the third set the function value should be at most or the bound should be from the lower side. So, one is from the higher side, one has to between a bandwidth and another case it has to be from the lower side.

So, what you are trying to do, if you consider this some of the problems, some of the problems what we have I will only write the functional form of g's are greater than type. I am writing b in a very general sense. Some of them of the less than type and some of them can be equal, but this equality I am basically putting it in the sense that there is a bound for that.

So, for the greater than type, it will be b 2 plus infinity. For the case of the bounded one, I will basically have some b 1 into b 2. So, this b 1 b 2 would change accordingly and other case it will be minus infinity and b. So, you have already divided the whole set of three regions upper bound, lower bound and some bound in between based on which you will assign your values.

(Refer Slide Time: 22:44)



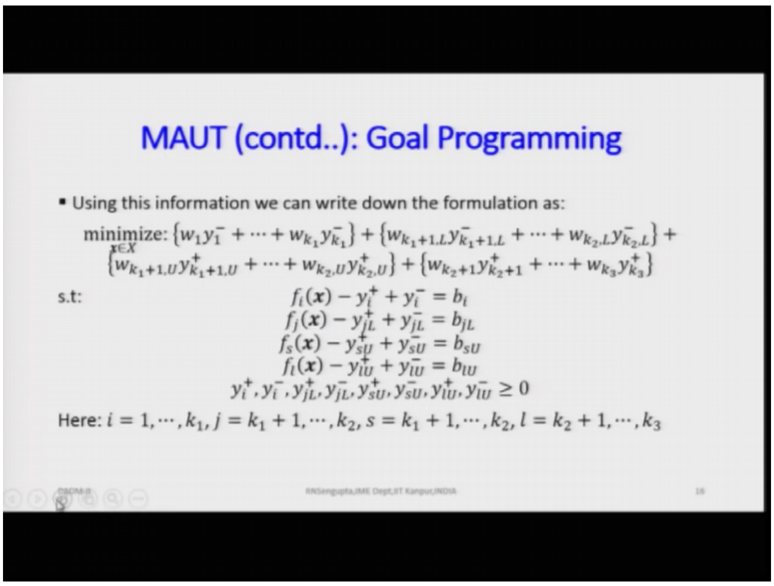
So, remember for any deviation from restrictions imposed they would be some functional weightages. So, if you remember the weightages which I had given. I had given weightages or I would not say I, but generally the problem formulation has been done in such a way. So, the weightages are for the first set would be w 1 to w k 1. So, these are greater than type. For other set which are in between you will basically put an upper bound and a lower bound because as you putting an upper bound and a lower bound for the b's also; similarly you will put an upper bound and the lower bound for the w's also for the next second set. And, for the third set they would basically be some w's based on which you will put the lower bound.

So, what I am trying to do is these are the bounds for greater than. These are the bounds for in between upper and lower values and these are the bounds for the less than type. Now, this is basis basically make some sense based on the fact that if you have taken the course t q m 1 and t q m 2, it basically gives you the upper control limit and the lower limit in some sense it is like this. So, what you have is this. So, consider I am not marking anything on the x axis and y axis, it is not very important from the problem formulation, but to understand it.

So, you will basically have three regions and mark these regions with this colour. So, this is the upper bound which you have, this is the lower bound which you have. So, what you want to do is one set would be greater. So, they may come here, one set is in between. So, this dispersion over and below the main line is not equal need not be. So, this is for the second set, this is for the less than type. So, this will come here and these perturbations. If I consider the normal distribution where the area on to the left or the right of the mean would; depend on the level of confidence.

So, if I have 2 sigma, plus sigma and minus sigma would give me the level of confidence as say for example, above 67 percent plus 2 sigma, minus 2 sigma is the total width would be plus 4 sigma would give me an level of confidence total area about 95 percent. And if I have plus 3 sigma, minus 3 sigma it will give me the level of confidence about 99 percent. So, using this I have base basically been able to formulate I divided into three regions, better greater of attaining the attribute in between I am satisfied, over and other I am also satisfied and less than I am satisfied depending on the level of confidence I have for attaining that.

(Refer Slide Time: 26:06)



So, I club them into three regions. And this I have already been talking about. So, in the first set these are the values for which I want to basically minimize or maximize for the upper limits. These are the value or let me use a other colour so it will be easier for me. These are the values for which I do it between bounds and these are the values for which I do for above. So, based on that I formulate and I go for the solution method.

Solution method I am not discussing I am giving you the idea as I did for the reliability part and this part. So, with this I will end the second lecture in the 10th week and continue more discussions accordingly for the simple concepts. So, again I am mentioning these are a little bit difficult topics. We will only go through the simple problems for the assignments.

So, with this I will end the class, have a nice day and thank you very much for your attention.